Mean Delay Comparison among Multilevel Processor-Sharing Scheduling Disciplines

S. Aalto and U. Ayesta

REPORT No. 11, 2004/2005, fall

ISSN 1103-467X
ISRN IML-R- -11-04/05- -SE+fall

INSTITUT MITTAG-LEFFLER
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Mean delay comparison among multilevel processor-sharing scheduling disciplines

Samuli Aalto
Networking Laboratory, Helsinki University of Technology

Urtzi Ayesta
France Telecom R & D and INRIA Sophia Antipolis

Abstract

Multilevel Processor-Sharing (MLPS) scheduling disciplines permit to model a wide variety of non-anticipating scheduling disciplines. Such disciplines have recently attracted attention in the context of the Internet as an appropriate flow-level model for the bandwidth sharing obtained when priority is given to short TCP connections. In this article we compare the mean delay in an M/G/1 queue among MLPS disciplines when the hazard rate of the service time distribution is decreasing. The internal disciplines within levels may vary in the family $\{\text{FB}, \text{PS}, \text{FCFS}\}$. Our main result states that, given an MLPS discipline, the mean delay is reduced whenever a level is added by splitting an existing one. Furthermore, we show that given an MLPS discipline, the mean delay is reduced when an internal discipline is changed from FCFS to PS (or from PS to FB). For service time distributions with increasing and bounded hazard rate, we obtain the corresponding inverse results. By numerical means we quantify the reduction on the mean delay after the addition of levels. We demonstrate that the mean delay of an MLPS discipline can get close to the minimum feasible delay (FB) with just a few levels. This supports the claim that in order to improve the flow-level performance in the Internet, a simple two-class classification (mice and elephants) might be sufficient.

Keywords Scheduling, MLPS, PS, FB, LAS, FCFS, M/G/1, mean delay, unfinished truncated work

1 Introduction

Multilevel Processor-Sharing (MLPS) scheduling disciplines introduced by L. Kleinrock in the early 1970's [1] permit to model a wide variety of non-anticipating scheduling disciplines. A discipline is non-anticipating when the scheduler does not know the remaining service time of jobs. Such disciplines have recently attracted attention in the context of the Internet as an appropriate flow-level model for the bandwidth sharing obtained when priority is given to short TCP connections [2, 3, 4, 5].

An MLPS scheduling discipline π is defined by a finite set of level thresholds $a_1 < \cdots < a_N$ defining $N+1$ levels, $N \geq 0$. A job belongs to level n if its attained service is at least a_{n-1} but less than a_n, where $a_0 = 0$ and $a_{N+1} = \infty$. Between these levels, a strict priority discipline is applied with the lowest level having the highest priority. Thus, those jobs with attained service less than a_1 are served first. Within each level n, an internal discipline D_n^π is applied. The internal disciplines may vary in the set $\{\mathrm{FB}, \mathrm{PS}, \mathrm{FCFS}\}$, where FB refers to the Foreground-Background discipline, which gives priority to the job with the least attained service, PS to the Processor-Sharing discipline, which shares the service capacity evenly among all jobs, and FCFS to the ordinary First-Come-First-Served discipline. The FB discipline is also known as LAS (Least-Attained-Service).

Yashkov [6] has proven that FB minimizes the mean delay, i.e., the mean sojourn time, among work-conserving and non-anticipating disciplines whenever the service time distribution is of type DHR (Decreasing Hazard Rate). Righter and Shanthikumar [7] proved that, under the DHR condition, FB minimizes the queue length even stochastically.

The PS discipline has been proposed as an appropriate model for the bandwidth sharing among (non-prioritized) TCP flows in a bottleneck router [8, 9, 10, 11]. According to a similar reasoning, PS is a relevant internal discipline for the MLPS disciplines that model the bandwidth sharing among prioritized TCP flows [2, 4]. On the other hand, flow sizes in the Internet have been modelled by, e.g., Pareto and hyperexponential distributions [12, 13]. The latter type satisfies the DHR condition, while the Pareto distribution defined in [12] by

$$P\{X \leq x\} = 1 - \left(\frac{k}{x}\right)^\alpha, \qquad x \geq k,$$

has a decreasing hazard rate only from the lower limit k on. However, if the Pareto distribution is defined as in [13] by

$$P\{X \leq x\} = 1 - \left(\frac{1}{1+cx}\right)^\alpha, \qquad x \geq 0,$$

the DHR condition is satisfied.

The following results have been derived for the MLPS disciplines whose internal disciplines vary in the set $\{\mathrm{FB}, \mathrm{PS}\}$ (with FCFS excluded) in the context of M/G/1 queues. Aalto et al. [14] proved that such MLPS disciplines with just *two* levels are better than PS with respect to the mean delay whenever the hazard rate of the service time distribution is decreasing, and vice versa if the hazard rate is increasing and bounded. In [15], the same authors compared, in a similar way, an MLPS discipline with *any* number of levels to the PS discipline.

In this article we compare the mean delay in an M/G/1 queue among MLPS disciplines. The internal disciplines may vary in the original set $\{\mathrm{FB}, \mathrm{PS}, \mathrm{FCFS}\}$. From the previous analysis an important question remained widely open: Given two MLPS disciplines, which one is better in the mean delay sense? Our main result states that given an MLPS discipline, the mean delay is reduced under the DHR condition if a level is added by splitting an existing one. As the number of levels grows to infinity, the MLPS discipline mimics the behaviour of an FB queue. Thus, our result provides a constructive way to demonstrate the optimality of FB. Furthermore, we show that given an MLPS discipline, the mean delay is reduced if an internal discipline is changed from FCFS to PS (or from PS to FB) and the hazard rate of the service time distribution is decreasing. These two results define a natural partial order among the MLPS disciplines so that we may answer to the question posed above if one MLPS discipline is derived from another by splitting levels and/or changing internal disciplines from FCFS to PS (or from PS to FB). For service time distributions with increasing and bounded hazard rate, we obtain the corresponding inverse results. By numerical means we quantify the reduction on the mean delay after the addition of levels. We show that the mean delay of an MLPS discipline can get close to the minimum feasible delay (FB) with just a few levels.

The rest of the paper is organized as follows. In Section 2 we explain the notation used throughout the paper. In Section 2 we also present our main new results on the mean delay comparison among MLPS disciplines. In Section 3 we analyze the effect of changing an internal discipline, and in Section 4 we consider the effect of splitting a level for an MLPS discipline. In Section 5 we give illustrative numerical examples and study by numerical means the mean delay of MLPS disciplines. Section 6 concludes the paper

2 Notation and results

In this section we present the notation used and the results of the paper. The key variable is unfinished truncated work. The fundamental relation between the mean delay and the mean unfinished truncated work is presented in (5). Some of the results are valid for each sample path, while the others are related to mean values only. The main new results are given in Theorems 2 and 3 at the end of this section.

We denote by MLPS the family of MLPS disciplines π for which $D_n^\pi \in \{\text{FB}, \text{PS}, \text{FCFS}\}$ for all n, and by MLPS* the family of MLPS disciplines π for which $D_n^\pi \in \{\text{FB}, \text{PS}\}$ for all n. Furthermore, we denote by $(N+1)$PS the family of MLPS disciplines π with $N+1$ levels for which $D_n^\pi = \text{PS}$ for all n. Thus, for example, 1PS refers to the PS discipline alone, 2PS to the PS+PS disciplines, 3PS to the PS+PS+PS disciplines etc. All these disciplines belong to the class Π of work-conserving and non-anticipating disciplines, which do not idle when there are jobs waiting and neither use they any information about the remaining service times of jobs. Among the disciplines $\{\text{FB}, \text{PS}, \text{FCFS}\}$, we define the following order relation:

$$\text{FB} \prec \text{PS} \prec \text{FCFS}.$$

In addition, among these disciplines, we denote $D \preceq D'$ if and only if $D = D'$ or $D \prec D'$.

2.1 Sample path results

Consider a single server queueing system starting empty at time $t = 0$ and obeying a scheduling discipline π. We assume that the jobs arrive one at a time.

Unfinished truncated work

Let A_i denote the arrival time of job i, S_i its service time, and $X_i^\pi(t)$ its attained service at time t. Let $\mathcal{A}(t)$ denote the set of jobs arrived until time t,

$$\mathcal{A}(t) = \{i : A_i \leq t\},$$

and $A(t) = |\mathcal{A}(t)|$. Let $\mathcal{N}^\pi(t)$ denote the set of jobs in the system at time t,

$$\mathcal{N}^\pi(t) = \{i \in \mathcal{A}(t) : X_i^\pi(t) < S_i\},$$

and $N^\pi(t) = |\mathcal{N}^\pi(t)|$. Furthermore, for all $x \geq 0$, let $\mathcal{N}_x^\pi(t)$ denote the set of those jobs in the system whose attained service is less than x,

$$\mathcal{N}_x^\pi(t) = \{i \in \mathcal{A}(t) : X_i^\pi(t) < (S_i \wedge x)\},$$

where $(S_i \wedge x) = \min\{S_i, x\}$, and $N_x^\pi(t) = |\mathcal{N}_x^\pi(t)|$. Let $U_x^\pi(t)$ denote the unfinished truncated work with truncation threshold x at time t,

$$U_x^\pi(t) = \sum_{i \in \mathcal{N}_x^\pi(t)} ((S_i \wedge x) - X_i^\pi(t)). \tag{1}$$

An alternative expression for the unfinished truncated work is as follows [14]:

$$U_x^\pi(t) = \sum_{i=1}^{A(t)} (S_i \wedge x) - \int_0^t \sigma_x^\pi(u)\, du, \tag{2}$$

where $\sigma_x^\pi(t)$ refers to the total rate at which the jobs with attained service less than x are served at time t. The limiting value $U_\infty^\pi(t)$ is the ordinary unfinished work, which is the same for all work conserving disciplines.

Known results

Proposition 1 *[14, Prop. 5] Let π be any scheduling discipline. Then $U_x^{\mathrm{FB}}(t) \leq U_x^\pi(t)$ for all $x \geq 0$ and $t \geq 0$.*

In [14], the authors restricted themselves to the class Π of work-conserving and non-anticipating disciplines. However, these properties are not at all utilized in the proof presented there. Therefore, π can really be *any* discipline.

It is also easy to see that, for any $\pi \in$ MLPS with a_n as a level threshold, $U_{a_n}^\pi(t) = U_{a_n}^{\mathrm{FB}}(t)$ for all $t \geq 0$.

Proposition 2 *[14, Prop. 8(i)] Let $\pi \in$ MLPS with thresholds $\{a_1, \ldots, a_N\}$ and $\pi' \in$ MLPS with thresholds $\{a'_1, \ldots, a'_{N'}\}$. Assume that there exist $n \in \{1, \ldots, N+1\}$ and $n' \in \{1, \ldots, N'+1\}$ such that $a_{n-1} = a'_{n'-1}$, $a_n = a'_{n'}$, and $D_n^\pi = D_{n'}^{\pi'}$. Then $U_x^\pi(t) = U_x^{\pi'}(t)$ for all $a_{n-1} \leq x \leq a_n$ and $t \geq 0$.*

This is a fundamental locality result for MLPS disciplines saying that the unfinished truncated work within a level depends only on the internal discipline of that level (but not on the other internal disciplines). In [14], the result is formulated only for class MLPS*, but the proof presented there is clearly valid for all MLPS disciplines. Thus, if we are only interested in the values of the unfinished truncated work $U_x^\pi(t)$ at level $a_{n-1} \leq x \leq a_n$, we may replace the original MLPS discipline π by a three-level discipline π' with thresholds $\{a_{n-1}, a_n\}$ and internal discipline $D_2^{\pi'} = D_n^\pi$. For $n = 1$, this simplifies to a two-level discipline.

New results

Proposition 3 *Let $\pi, \pi' \in$ MLPS with thresholds $\{a_1, \ldots, a_N\}$ such that $D_n^\pi \preceq D_n^{\pi'}$ for all $n \in \{1, \ldots, N+1\}$. Then $U_x^\pi(t) \leq U_x^{\pi'}(t)$ for all $x \geq 0$ and $t \geq 0$.*

This partially new result, generalizing [14, Prop. 8(ii)] from class MLPS* to class MLPS, is the first step in ordering the MLPS disciplines. It allows us to compare different MLPS disciplines with the same thresholds. The proof will be given in Section 3.

Note further that, by changing all the internal disciplines to FB, the unfinished truncated work is reduced. On the other hand, an MLPS discipline π for which $D_n^\pi =$ FB for all n is nothing but the ordinary FB discipline [1], which is in line with Proposition 1.

2.2 Mean value results

Consider an M/G/1 queue obeying a scheduling discipline $\pi \in \Pi$. Let λ denote the arrival rate and S the service time of a job. We assume that $E[S] < \infty$ and $\rho = \lambda E[S] < 1$. Furthermore, we assume that the service time distribution is continuous with the corresponding density function denoted by $f(x)$. Let $F(x) = \int_0^x f(y)\,dy$ and $\overline{F}(x) = 1 - F(x)$. The corresponding hazard rate function is denoted by $h(x) = f(x)/\overline{F}(x)$.

Mean unfinished truncated work

Let U_x^π denote the unfinished truncated work with truncation threshold x and $T^\pi(y)$ the delay of a job with service time of y time units. By [1, Eq. (4.60)],

$$\overline{U}_x^\pi = \lambda \int_0^x \overline{T}^\pi(y)\overline{F}(y)\,dy, \tag{3}$$

where $\overline{U}_x^\pi = E[U_x^\pi]$ and $\overline{T}^\pi(y) = E[T^\pi(y)]$. Thus,

$$\frac{\partial}{\partial x}\overline{U}_x^\pi = \lambda\, \overline{T}^\pi(x)\, \overline{F}(x). \tag{4}$$

Let then T^π denote the delay of any job. It follows from (4) that

$$\overline{T}^\pi = \frac{1}{\lambda}\int_0^\infty (\overline{U}_x^\pi)' h(x)\,dx, \tag{5}$$

where $\overline{T}^\pi = E[T^\pi]$ and $(\overline{U}_x^\pi)' = \frac{\partial}{\partial x}\overline{U}_x^\pi$.

Known results

Proposition 4 *[14, Props. 1 and 2] Let $\pi, \pi' \in \Pi$ such that $\overline{U}_x^{\pi} \le \overline{U}_x^{\pi'}$ for all $x \ge 0$.*

(i) If the hazard rate $h(x)$ is decreasing, then $\overline{T}^{\pi} \le \overline{T}^{\pi'}$.

(ii) If the hazard rate $h(x)$ is increasing and bounded, then $\overline{T}^{\pi} \ge \overline{T}^{\pi'}$.

This is a key result that allows us to compare the mean delay of two scheduling disciplines belonging to class Π through comparison of the mean unfinished truncated work.

Proposition 5 *Let $\pi \in \Pi$. Then $\overline{U}_x^{\pi} \le \overline{U}_x^{\mathrm{FCFS}}$ for all $x \ge 0$.*

This result complements the earlier result concerning the optimality of FB with respect to the unfinished truncated work. The proof will be given in Subsection 4.3. A similar result, however, with different arguments used for the proof, appeared in [3, Lemma 2.4].

Theorem 1 *Let $\pi \in \Pi$.*

(i) If the hazard rate $h(x)$ is decreasing, then $\overline{T}^{\mathrm{FB}} \le \overline{T}^{\pi} \le \overline{T}^{\mathrm{FCFS}}$.

(ii) If the hazard rate $h(x)$ is increasing and bounded, then $\overline{T}^{\mathrm{FB}} \ge \overline{T}^{\pi} \ge \overline{T}^{\mathrm{FCFS}}$.

This is an immediate consequence of Propositions 1, 4, and 5. The result itself is not at all new, see, e.g., [6, 16, 17]. This is just a novel way to get it.

New results

Proposition 6 *Let $N \ge 1$, $n \in \{1, \dots, N\}$, $\pi \in$ MLPS with thresholds $\{a_1, \dots, a_N\}$, and $\pi' \in$ MLPS with thresholds $\{a_1, \dots, a_{n-1}, a_{n+1}, \dots, a_N\}$ such that $D_n^{\pi'} = D_n^{\pi} = D_{n+1}^{\pi} = \mathrm{FCFS}$ and*

$$D_m^{\pi'} = \begin{cases} D_m^{\pi} & m = 1, \dots, n-1, \\ D_{m+1}^{\pi} & m = n+1, \dots, N. \end{cases}$$

Then, $\overline{U}_x^{\pi} \le \overline{U}_x^{\pi'}$ for all $x \ge 0$.

This new result shows that, with respect to the mean unfinished truncated work, it is beneficial to split one level with FCFS as an internal discipline to two contiguous levels with the FCFS discipline applied at both levels. The proof will be given in Subsection 4.4.

Proposition 7 *Let $N \ge 1$, $n \in \{1, \dots, N\}$, $\pi \in$ MLPS with thresholds $\{a_1, \dots, a_N\}$, and $\pi' \in$ MLPS with thresholds $\{a_1, \dots, a_{n-1}, a_{n+1}, \dots, a_N\}$ such that $D_n^{\pi'} = D_n^{\pi} = D_{n+1}^{\pi} = \mathrm{PS}$ and*

$$D_m^{\pi'} = \begin{cases} D_m^{\pi} & m = 1, \dots, n-1, \\ D_{m+1}^{\pi} & m = n+1, \dots, N. \end{cases}$$

Then, $\overline{U}_x^{\pi} \le \overline{U}_x^{\pi'}$ for all $x \ge 0$.

This is a counterpart of the previous result concerning a level with PS as an internal discipline instead of FCFS. The proof will be given in Subsection 4.5.

Theorem 2 *Let* $\pi, \pi' \in \text{MLPS}$ *with thresholds* $\{a_1, \ldots, a_N\}$ *such that* $D_n^\pi \preceq D_n^{\pi'}$ *for all* $n \in \{1, \ldots, N+1\}$.

(i) If the hazard rate $h(x)$ *is decreasing, then* $\overline{T}^\pi \leq \overline{T}^{\pi'}$.

(ii) If the hazard rate $h(x)$ *is increasing and bounded, then* $\overline{T}^\pi \geq \overline{T}^{\pi'}$.

This new result is an immediate consequence of Propositions 3 and 4.

Theorem 3 *Let* $N \geq 1$, $n \in \{1, \ldots, N\}$, $\pi \in \text{MLPS}$ *with thresholds* $\{a_1, \ldots, a_N\}$, *and* $\pi' \in \text{MLPS}$ *with thresholds* $\{a_1, \ldots, a_{n-1}, a_{n+1}, \ldots, a_N\}$ *such that* $D_n^\pi = D_{n+1}^\pi = D_n^{\pi'}$ *and*

$$D_m^{\pi'} = \begin{cases} D_m^\pi & m = 1, \ldots, n-1, \\ D_{m+1}^\pi & m = n+1, \ldots, N. \end{cases}$$

(i) If the hazard rate $h(x)$ *is decreasing, then* $\overline{T}^\pi \leq \overline{T}^{\pi'}$.

(ii) If the hazard rate $h(x)$ *is increasing and bounded, then* $\overline{T}^\pi \geq \overline{T}^{\pi'}$.

This new result is an immediate consequence of Propositions 4, 6 and 7. Note that splitting an FB level into two FB levels does not change anything.

3 Changing internal disciplines

In this section we prove the sample path result presented in Proposition 3. Thus, we consider a single server queueing system starting empty at time $t = 0$ and assume that the jobs arrive one at a time. We start with a known result for MLPS* disciplines taken from [14].

Proposition 8 *[14, Prop. 8(ii)] Let* $\pi, \pi' \in \text{MLPS}^*$ *with thresholds* $\{a_1, \ldots, a_N\}$ *such that* $D_n^\pi \preceq D_n^{\pi'}$ *for all* $n \in \{1, \ldots, N+1\}$. *Then* $U_x^\pi(t) \leq U_x^{\pi'}(t)$ *for all* $x \geq 0$ *and* $t \geq 0$.

We present now a new auxillary result that allows us to generalize the previous result to all MLPS disciplines.

Proposition 9 *Let* $\pi \in \text{MLPS}$ *with thresholds* $\{a_1, \ldots, a_N\}$ *and* $\pi' \in \text{MLPS}$ *with thresholds* $\{a'_1, \ldots, a'_{N'}\}$. *Assume that there exist* $n \in \{1, \ldots, N+1\}$ *and* $n' \in \{1, \ldots, N'+1\}$ *such that* $a_{n-1} = a'_{n'-1}$, $a_n = a'_{n'}$, $D_n^\pi = \text{PS}$, *and* $D_{n'}^{\pi'} = \text{FCFS}$. *Then* $U_x^\pi(t) \leq U_x^{\pi'}(t)$ *for all* $a_{n-1} \leq x \leq a_n$ *and* $t \geq 0$.

Proof. Let $a_{n-1} \le x \le a_n$. We prove the claim by induction with respect to arrival epochs A_k.

1° During the interval $[0, A_1)$ both systems are empty. Thus the claim is trivially true for all $t < A_1$.

2° Let $k \in \{1, 2, \ldots\}$, and assume that the claim is true for all $t < A_k$. We will show that it is also true in the interval $[A_k, A_{k+1})$.

We divide the interval $[A_k, A_{k+1})$ into three consecutive periods I_1, I_2, and I_3, with the following starting (b) and ending (e) points:

$$\begin{array}{ll} I_1^{\mathrm{b}} = A_k, & I_1^{\mathrm{e}} = \sup\{I_1^{\mathrm{b}} < t \le A_{k+1} \mid N_{a_{n-1}}^{\pi}(t) > 0\}, \\ I_2^{\mathrm{b}} = I_1^{\mathrm{e}}, & I_2^{\mathrm{e}} = \sup\{I_2^{\mathrm{b}} < t \le A_{k+1} \mid N_{a_n}^{\pi}(t) > 0\}, \\ I_3^{\mathrm{b}} = I_2^{\mathrm{e}}, & I_3^{\mathrm{e}} = A_{k+1}. \end{array}$$

During interval I_1 both systems give service only to those customers whose attained service is less than a_{n-1}. During interval I_2 there are not any longer such customers in either system, and both systems give service to those customers whose attained service is at least a_{n-1} but less than a_n. Finally, in interval I_3, there are not any longer jobs with attained service less than a_n in either system. Note that I_1 is always of positive length, whereas I_2 and I_3 may vanish. The three intervals I_1 – I_3 are considered in 2.1° – 2.3°, respectively.

2.1° Consider the interval I_1. Since $x \ge a_{n-1}$ and, during this interval, strict priority is given (in both systems) to those customers with attained service time less than a_{n-1}, we have, for all $t \in I_1$,

$$\begin{aligned} U_x^{\pi}(t) &= U_x^{\pi}((A_k)^-) + (S_k \wedge x) - t + A_k \\ &\le U_x^{\pi'}((A_k)^-) + (S_k \wedge x) - t + A_k \\ &= U_x^{\pi'}(t), \end{aligned}$$

where the inequality is due to the induction assumption. This is enough if the interval I_1 ends at time A_{k+1} when a new customer arrives. Otherwise we have to consider the interval I_2, too.

2.2° Consider the interval I_2. During this interval, strict priority is given (in both systems) to those customers with attained service time at least a_{n-1} but less than a_n. Let $t \in I_2$ and denote $\mathcal{M}(t) = \mathcal{N}_{a_n}^{\pi}(t) \cup \mathcal{N}_{a_n}^{\pi'}(t)$. Thus, $\mathcal{M}(t)$ comprises of the customers that are priority customers, at least, in one of the systems at time t. Since a_{n-1} and a_n are level thresholds for both π and π', we have, for all $i \in \mathcal{M}(t)$,

$$a_{n-1} \le X_i^{\pi}(t) \le a_n, \quad a_{n-1} \le X_i^{\pi'}(t) \le a_n.$$

Now,

$$U_x^{\pi}(t) = \sum_{i \in \mathcal{M}(t)} (S_i \wedge x) - \sum_{i \in \mathcal{M}(t)} (X_i^{\pi}(t) \wedge x), \tag{6}$$

$$U_x^{\pi'}(t) = \sum_{i \in \mathcal{M}(t)} (S_i \wedge x) - \sum_{i \in \mathcal{M}(t)} (X_i^{\pi'}(t) \wedge x). \tag{7}$$

Note that the first sum is the same for both disciplines. For $x = a_n$, we get

$$\begin{aligned}
U^{\pi}_{a_n}(t) &= \sum_{i\in\mathcal{M}(t)} (S_i \wedge a_n) - \sum_{i\in\mathcal{M}(t)} (X^{\pi}_i(t) \wedge a_n) \\
&= \sum_{i\in\mathcal{M}(t)} (S_i \wedge a_n) - \sum_{i\in\mathcal{M}(t)} X^{\pi}_i(t), \\
U^{\pi'}_{a_n}(t) &= \sum_{i\in\mathcal{M}(t)} (S_i \wedge a_n) - \sum_{i\in\mathcal{M}(t)} (X^{\pi'}_i(t) \wedge a_n) \\
&= \sum_{i\in\mathcal{M}(t)} (S_i \wedge a_n) - \sum_{i\in\mathcal{M}(t)} X^{\pi'}_i(t).
\end{aligned}$$

Since a_n is a level threshold for both disciplines, the unfinished truncated work with truncation threshold a_n is the same in both systems. Thus,

$$\sum_{i\in\mathcal{M}(t)} X^{\pi}_i(t) = \sum_{i\in\mathcal{M}(t)} X^{\pi'}_i(t). \tag{8}$$

Let then $I(t)$ denote the index of the customer who is served in the π' system at time t. Because of the internal FCFS discipline of the π' system, we observe that, for all $i \in \mathcal{M}(t)$ such that $i < I(t)$,

$$X^{\pi}_i(t) < (S_i \wedge a_n) = X^{\pi'}_i(t), \tag{9}$$

and, for all $i \in \mathcal{M}(t)$ such that $i > I(t)$,

$$X^{\pi}_i(t) \geq a_{n-1} = X^{\pi'}_i(t). \tag{10}$$

Furthermore, due to the internal PS discipline in the π system, we have, for all $i, j \in \mathcal{M}(t)$ such that $i < j$,

$$X^{\pi}_i(t) \geq X^{\pi}_j(t). \tag{11}$$

The next step is to utilize Lemma 2 presented in Appendix A. Consider customers $i \in \mathcal{M}(t)$. Let $N = |\mathcal{M}(t)|$, and reindex these customers in their arrival order from 1 to N. Let $r(i)$ denote the new index of customer i. Furthermore, let $a_{r(i)} = X^{\pi}_i(t)$ and $b_{r(i)} = X^{\pi'}_i(t)$ for all $i \in \mathcal{M}(t)$. In addition, let

$$m = \begin{cases} r(I(t)), & \text{if } X^{\pi}_{I(t)}(t) \leq X^{\pi'}_{I(t)}(t), \\ r(I(t)) - 1, & \text{if } X^{\pi}_{I(t)}(t) > X^{\pi'}_{I(t)}(t). \end{cases}$$

With the results given above in (8), (9), (10), and (11), it is easy to check that all the assumptions of Lemma 2 are valid with these choices. Now, by applying Lemma 2, we conclude that

$$\sum_{i\in\mathcal{M}(t)} (X^{\pi}_i(t) \wedge x) \geq \sum_{i\in\mathcal{M}(t)} (X^{\pi'}_i(t) \wedge x).$$

Thus, by (6) and (7), we have, for any $t \in I_2$,

$$U^{\pi}_x(t) \leq U^{\pi'}_x(t).$$

This is enough if the interval I_2 ends at time A_{k+1} when a new job arrives. Otherwise we have to consider the final interval I_3, too.

2.3° Consider finally the interval I_3, in which there are not any more customers with attained service less than a_n in either system. Since $x \leq a_n$, we deduce that, for all $t \in I_3$,

$$U_x^\pi(t) = U_x^{\pi'}(t) = 0.$$

This completes the proof. □

Proposition 3 follows now straightforwardly from Propositions 2, 8, and 9.

4 Splitting levels

In this section we consider the effect of splitting levels. We start with some sample path results, and then present the proofs of Propositions 6 and 7 separately. As a byproduct, we get a proof for the extremality result of FCFS presented in Proposition 5.

4.1 Sample path results

We consider a single server queueing system starting empty at time $t = 0$ and assume that the jobs arrive one at a time. We start with a known result for $(N+1)$PS disciplines taken from [15].

Proposition 10 *[15, Prop. 6] Let $N \geq 1$, $\pi \in (N+1)$PS with thresholds $\{a_1, \ldots, a_N\}$, and $\pi' \in N$PS with thresholds $\{a_1, \ldots, a_{N-1}\}$. Then $U_x^\pi(t) \leq U_x^{\pi'}(t)$ for all $x \leq a_N$ and $t \geq 0$.*

Let then $U_x^\pi(t; S \wedge b)$ denote the unfinished truncated work in a system where the original service times $S_1, S_2, \ldots$ are replaced by their truncated versions $S_1 \wedge b, S_2 \wedge b, \ldots$.

Proposition 11 *Let $N \geq 1$, $n \in \{1, \ldots, N\}$, and $\pi \in$ MLPS with thresholds $\{a_1, \ldots, a_N\}$. Then, $U_x^\pi(t) = U_x^\pi(t; S \wedge a_n)$ for all $x \leq a_n$ and $t \geq 0$.*

Proof. Let $x \leq a_n$ and $t \geq 0$. Since the two systems follow the same rules as regards the jobs with attained service time less than a_n (which is a level threshold in the original system and the service time truncation threshold in the modified system), we surely have

$$\mathcal{N}_x^\pi(t) = \mathcal{N}_x^\pi(t; S \wedge a_n).$$

and, for all $i \in \mathcal{N}_x^\pi(t)$,

$$X_i^\pi(t) = X_i^\pi(t; S \wedge a_n).$$

The claim follows now from (1). □

4.2 Preliminaries for the mean value results

From this on, we consider an M/G/1 queue with $\rho < 1$. In this subsection we recall the properties of the conditional mean delay for MLPS disciplines.

As stated in Proposition 2, the unfinished truncated work within a level depends only on the internal discipline of that level (but not on the other internal disciplines). There is a similar result concerning the conditional mean delay. More precisely, it is known that, for all $a_{n-1} < x \le a_n$, the conditional mean delay is of the following form:

$$\overline{T}^{\pi}(x) = f_1(a_{n-1}) + f_2^{D_n^{\pi}}(a_{n-1}, x, a_n). \tag{12}$$

Thus, within a level, it depends only on the internal discipline of that level. The three possible internal disciplines are discussed separately below. Before that we, however, recall some elementary results for systems with truncated service times. All the results of this subsection may be found, e.g., in [1].

Truncated service times

Let $x \ge 0$, and replace, for a while, the service times S by their truncated versions $S \wedge x = \min\{S, x\}$. It is easy to see that

$$E[S \wedge x] = \int_0^x \overline{F}(z)\, dz, \quad E[(S \wedge x)^2] = 2\int_0^x z\overline{F}(z)\, dz. \tag{13}$$

Furthermore, let $\rho_x = \lambda E[(S \wedge x)]$ denote the truncated load. Clearly, $\rho_x \le \rho < 1$ for all x. For future purposes we also present the derivatives of these quantities with respect to the truncation threshold x:

$$\tfrac{\partial}{\partial x} E[S \wedge x] = \overline{F}(x), \quad \tfrac{\partial}{\partial x} E[(S \wedge x)^2] = 2x\overline{F}(x),$$
$$\tfrac{\partial}{\partial x} \rho_x = \lambda \overline{F}(x).$$

The mean workload (i.e., unfinished work) for a work conserving M/G/1 queue with truncated service times is, by the Pollaczek-Khinchin formula,

$$\overline{W}_x = \frac{\lambda E[(S \wedge x)^2]}{2(1-\rho_x)}. \tag{14}$$

Regarding the derivative, it is easy to verify that

$$\frac{\partial}{\partial x}\overline{W}_x = \lambda\, \frac{\overline{W}_x + x}{1-\rho_x}\, \overline{F}(x) = \lambda\, \overline{T}^{\mathrm{FB}}(x)\, \overline{F}(x). \tag{15}$$

The last equation follows from a result given below. It is interesting to compare this with the earlier result (4) concerning the derivative of the mean unfinished truncated work $\overline{U}_x^{\pi}$.

It is also worth to mention that the unfinished work W_x for an M/G/1 queue with truncated service times $(S \wedge x)$ differs, in general, from the unfinished truncated work U_x^{π} for an M/G/1 queue with original service times S. The former one is the same for all work-conserving disciplines, while the latter one depends on the discipline π used. Only for the FB discipline, $W_x = U_x^{\mathrm{FB}}$.

Of course, when $x \to \infty$, we get the ordinary Pollaczek-Khinchin formula,

$$\overline{W}_\infty = \frac{\lambda E[S^2]}{2(1-\rho)}.$$

Conditional mean delay for MLPS disciplines within an FB level

Let us return to the original service times S. First we consider the conditional mean delay for MLPS disciplines within an FB level. If there is just one level, then, for all $x \geq 0$,

$$\overline{T}^{\mathrm{FB}}(x) = \frac{\overline{W}_x + x}{1-\rho_x}. \tag{16}$$

Regarding the derivative, it is easy to verify that

$$(\overline{T}^{\mathrm{FB}})'(x) = \frac{1 + 2\lambda \overline{T}^{\mathrm{FB}}(x)\overline{F}(x)}{1-\rho_x}. \tag{17}$$

If π is an MLPS discipline with FB applied at level n, then, for all $a_{n-1} < x \leq a_n$,

$$\overline{T}^{\pi}(x) = \overline{T}^{\mathrm{FB}}(x) = \frac{\overline{W}_x + x}{1-\rho_x}. \tag{18}$$

Conditional mean delay for MLPS disciplines within an FCFS level

Next we consider the conditional mean delay for MLPS disciplines within an FCFS level. If there is just one level, then, for all $x \geq 0$,

$$\overline{T}^{\mathrm{FCFS}}(x) = \overline{W}_\infty + x. \tag{19}$$

If π is an MLPS discipline with FCFS applied at level n, then, for all $a_{n-1} < x \leq a_n$,

$$\overline{T}^{\pi}(x) = \frac{\overline{W}_{a_n} + x}{1-\rho_{a_{n-1}}}. \tag{20}$$

Conditional mean delay for MLPS disciplines within a PS level

Finally we consider the conditional mean delay for MLPS disciplines within a PS level. If there is just one level, then, for all $x \geq 0$,

$$\overline{T}^{\mathrm{PS}}(x) = \frac{x}{1-\rho}. \tag{21}$$

If π is an MLPS discipline with PS applied at level n, then, for all $a_{n-1} < x \leq a_n$,

$$\overline{T}^{\pi}(x) = \overline{T}^{\mathrm{FB}}(a_{n-1}) + \frac{\alpha(x - a_{n-1})}{1-\rho_{a_{n-1}}}. \tag{22}$$

Here $\alpha(\cdot)$ refers to the conditional mean delay in an M^{X}/G/1 queue with batch arrivals and PS discipline, see Appendix B. The batches arrive with rate λ. The mean batch size is

$$\bar{a} = \frac{\overline{F}(a_{n-1})}{1-\rho_{a_{n-1}}},$$

and the parameter b, denoting the mean number of additional customers in an arriving batch where there is at least one customer, is given by

$$b = 2\lambda \overline{T}^{\mathrm{FB}}(a_{n-1})\overline{F}(a_{n-1}).$$

In addition, since the customers in this batch arrival queue are those in the original MLPS system whose service times are greater than the level threshold a_{n-1}, the tail distribution of the service time in the batch arrival queue is as follows:

$$\overline{B}(t) = \begin{cases} \dfrac{\overline{F}(a_{n-1}+t)}{\overline{F}(a_{n-1})}, & 0 \le t < a_n - a_{n-1}, \\ 0, & a_n - a_{n-1} \le t < \infty. \end{cases}$$

Thus, the integral equation (31) for the derivative $\alpha'(t)$ presented in Appendix B takes the following form:

$$\begin{aligned} \alpha'(t) \;&=\; 1 \;+\; 2\lambda \overline{T}^{\mathrm{FB}}(a_{n-1})\overline{F}(a_{n-1}+t) \\ &+\; \frac{\lambda}{1-\rho_{a_{n-1}}} \int_0^{a_n - a_{n-1} - t} \alpha'(u)\overline{F}(a_{n-1}+t+u)\,du \\ &+\; \frac{\lambda}{1-\rho_{a_{n-1}}} \int_0^{t} \alpha'(u)\overline{F}(a_{n-1}+t-u)\,du. \end{aligned} \tag{23}$$

By (22) and (23), the derivative of the conditional mean delay can be written, for all $a_{n-1} < x \le a_n$, as follows:

$$\begin{aligned} (\overline{T}^{\pi})'(x) \;&=\; \frac{1}{1-\rho_{a_{n-1}}} \Big(1 \;+\; 2\lambda \overline{T}^{\mathrm{FB}}(a_{n-1})\overline{F}(x) \\ &+\; \frac{\lambda}{1-\rho_{a_{n-1}}} \int_0^{a_n - x} \alpha'(u)\overline{F}(x+u)\,du \\ &+\; \frac{\lambda}{1-\rho_{a_{n-1}}} \int_0^{x - a_{n-1}} \alpha'(u)\overline{F}(x-u)\,du \Big). \end{aligned} \tag{24}$$

If $n = N+1$, so that PS is applied at the highest level, we have, for all $x > a_N$,

$$\begin{aligned} (\overline{T}^{\pi})'(x) \;&=\; \frac{1}{1-\rho_{a_N}} \Big(1 \;+\; 2\lambda \overline{T}^{\mathrm{FB}}(a_N)\overline{F}(x) \\ &+\; \frac{\lambda}{1-\rho_{a_N}} \int_0^{\infty} \alpha'(u)\overline{F}(x+u)\,du \\ &+\; \frac{\lambda}{1-\rho_{a_N}} \int_0^{x - a_N} \alpha'(u)\overline{F}(x-u)\,du \Big). \end{aligned} \tag{25}$$

4.3 Extremal property of FCFS

In this subsection we present the proof of Proposition 5. We start with a small auxillary result presented in the following lemma.

Lemma 1 *Let $\pi, \pi' \in \Pi$ and $a < b$. Assume that*

(i) $\overline{U}_a^\pi \le \overline{U}_a^{\pi'}$, $\overline{U}_b^\pi \le \overline{U}_b^{\pi'}$, *and*

(ii) $\overline{T}^\pi(x) - \overline{T}^{\pi'}(x)$ *is non-decreasing for all* $a < x \le b$.

Then $\overline{U}_x^\pi \le \overline{U}_x^{\pi'}$ *for all* $a \le x \le b$.

Proof. Assume that there is $a < x < b$ such that $\overline{U}_x^\pi > \overline{U}_x^{\pi'}$. Consider now what happens if $\overline{T}^\pi(x) \le \overline{T}^{\pi'}(x)$. Due to (ii), this would imply that $\overline{T}^\pi(y) \le \overline{T}^{\pi'}(y)$ for all $a < y < x$. Then, by (3),

$$
\begin{aligned}
\overline{U}_x^\pi &= \overline{U}_a^\pi + \lambda \int_a^x \overline{T}^\pi(y)\overline{F}(y)\,dy \\
&\le \overline{U}_a^{\pi'} + \lambda \int_a^x \overline{T}^{\pi'}(y)\overline{F}(y)\,dy \;=\; \overline{U}_x^{\pi'},
\end{aligned}
$$

which is impossible. Thus, we conclude that $\overline{T}^\pi(x) > \overline{T}^{\pi'}(x)$. Due to (ii), $\overline{T}^\pi(y) > \overline{T}^{\pi'}(y)$ for all $x < y < b$. Then, by (3),

$$
\begin{aligned}
\overline{U}_b^\pi &= \overline{U}_x^\pi + \lambda \int_x^b \overline{T}^\pi(y)\overline{F}(y)\,dy \\
&> \overline{U}_x^{\pi'} + \lambda \int_x^b \overline{T}^{\pi'}(y)\overline{F}(y)\,dy \;=\; \overline{U}_b^{\pi'},
\end{aligned}
$$

which is impossible. Thus, $\overline{U}_x^\pi \le \overline{U}_x^{\pi'}$ for all $a \le x \le b$. □

Let us now recall Proposition 5 and then prove it.

Proposition 5 *Let $\pi \in \Pi$. Then $\overline{U}_x^\pi \le \overline{U}_x^{\mathrm{FCFS}}$ for all $x \ge 0$.*

Proof. Let $\pi \in \Pi$. By definition, $\overline{U}_0^\pi = \overline{U}_0^{\mathrm{FCFS}} = 0$. In addition, since both π and FCFS are work-conserving, we have $\overline{U}_\infty^\pi = \overline{U}_\infty^{\mathrm{FCFS}} = \overline{W}_\infty$. By (19), we have $(\overline{T}^{\mathrm{FCFS}})'(x) = 1$ for all $x \ge 0$. Since the conditional mean delay cannot increase slower than this, the difference $\overline{T}^\pi(x) - \overline{T}^{\mathrm{FCFS}}(x)$ is non-decreasing for all $x \ge 0$. Thus, by Lemma 1, we conclude that the claim is true. □

4.4 Splitting an FCFS level

In this subsection we present the proof of Proposition 6.

Proposition 6 *Let $N \geq 1$, $n \in \{1, \ldots, N\}$, $\pi \in$ MLPS with thresholds $\{a_1, \ldots, a_N\}$, and $\pi' \in$ MLPS with thresholds $\{a_1, \ldots, a_{n-1}, a_{n+1}, \ldots, a_N\}$ such that $D_n^{\pi'} = D_n^{\pi} = D_{n+1}^{\pi} =$ FCFS and*

$$D_m^{\pi'} = \begin{cases} D_m^{\pi} & m = 1, \ldots, n-1, \\ D_{m+1}^{\pi} & m = n+1, \ldots, N. \end{cases}$$

Then, $\overline{U}_x^{\pi} \leq \overline{U}_x^{\pi'}$ for all $x \geq 0$.

Proof. By Proposition 2, we have $\overline{U}_x^{\pi} = \overline{U}_x^{\pi'}$ for all $x \leq a_{n-1}$ and $x \geq a_{n+1}$. In particular,

$$\overline{U}_{a_{n-1}}^{\pi} = \overline{U}_{a_{n-1}}^{\pi'}, \qquad \overline{U}_{a_{n+1}}^{\pi} = \overline{U}_{a_{n+1}}^{\pi'}. \tag{26}$$

Thus, it is sufficient to consider the interval $a_{n-1} < x < a_{n+1}$. By (20), we have

$$\overline{T}^{\pi}(x) = \begin{cases} \dfrac{\overline{W}_{a_n} + x}{1 - \rho_{a_{n-1}}}, & a_{n-1} < x \leq a_n, \\ \dfrac{\overline{W}_{a_{n+1}} + x}{1 - \rho_{a_n}}, & a_n < x \leq a_{n+1}, \end{cases}$$

and

$$\overline{T}^{\pi'}(x) = \frac{\overline{W}_{a_{n+1}} + x}{1 - \rho_{a_{n-1}}}, \qquad a_{n-1} < x \leq a_{n+1}.$$

It is easy to verify that the difference $\overline{T}^{\pi}(x) - \overline{T}^{\pi'}(x)$ is now non-decreasing for all $a_{n-1} < x \leq a_{n+1}$. Thus, by Lemma 1, we conclude that the claim is true. □

4.5 Splitting a PS level

In this subsection we present the proof of Proposition 7. We start with a key result presented below.

Proposition 12 *Let $\pi \in$ 2PS with threshold a and $\pi' \in$ 2PS with threshold a' such that $a \leq a'$. Then, for all $x > a'$,*

$$(\overline{T}^{\pi})'(x) \leq (\overline{T}^{\pi'})'(x).$$

Proof. Let us denote $\pi = 2\mathrm{PS}(a)$. By (25), we have, for all $x > a$,

$$\begin{aligned} (\overline{T}^{2\mathrm{PS}(a)})'(x) &= \frac{\alpha'(x-a)}{1-\rho_a} \\ &= \frac{1}{1-\rho_a}\Big(1 \,+\, 2\lambda \overline{T}^{\mathrm{FB}}(a)\overline{F}(x) \\ &\quad + \frac{\lambda}{1-\rho_a}\int_0^{\infty} \alpha'(u)\overline{F}(x+u)\,du \\ &\quad + \frac{\lambda}{1-\rho_a}\int_0^{x-a} \alpha'(u)\overline{F}(x-u)\,du\Big). \end{aligned}$$

It is important to note that the function $\alpha'(\cdot)$ itself depends on a but not on x.

Let us then fix x and consider the derivative above as a function of the level threshold a,

$$g_x(a) \stackrel{\text{def}}{=} (\overline{T}^{2\text{PS}(a)})'(x) = \frac{\alpha'(x-a)}{1-\rho_a}, \qquad 0 \le a < x.$$

Then we get, for all $0 \le a < x$,

$$\begin{aligned}
(g_x)'(a) &= \frac{\partial}{\partial a}\frac{\alpha'(x-a)}{1-\rho_a} \\
&= \frac{1}{1-\rho_a}\left(g_x(a)\frac{\partial}{\partial a}\rho_a \;+\; \frac{\partial}{\partial a}\alpha'(x-a)\right) \\
&= \frac{1}{1-\rho_a}\left(\lambda g_x(a)\overline{F}(a) \;+\; 2\lambda\overline{F}(x)\frac{\partial}{\partial a}\overline{T}^{\text{FB}}(a)\right. \\
&\quad + \lambda\int_0^\infty \frac{\partial}{\partial a}\frac{\alpha'(u)}{1-\rho_a}\overline{F}(x+u)\,du \\
&\quad \left. + \;\lambda\int_0^{x-a} \frac{\partial}{\partial a}\frac{\alpha'(u)}{1-\rho_a}\overline{F}(x-u)\,du \;-\; \lambda g_x(a)\overline{F}(a)\right) \\
&= \frac{\lambda}{1-\rho_a}\left(2\frac{\partial}{\partial a}\overline{T}^{\text{FB}}(a)\overline{F}(x)\right. \\
&\quad + \int_0^\infty \frac{\partial}{\partial a}\frac{\alpha'(u)}{1-\rho_a}\overline{F}(x+u)\,du \\
&\quad \left. + \int_0^{x-a} \frac{\partial}{\partial a}\frac{\alpha'(u)}{1-\rho_a}\overline{F}(x-u)\,du\right).
\end{aligned}$$

Let us now fix a, for a while, and consider the following function:

$$\beta(z) \stackrel{\text{def}}{=} \frac{\partial}{\partial a}\frac{\alpha'(z)}{1-\rho_a}, \qquad z \ge 0.$$

It is crucial to observe that $(g_x)'(a) = \beta(x-a)$ for all $x > a$. The previous equation, which is also true for all $x > a$, implies that, for all $z \ge 0$ (case $z = 0$ by a continuity argument),

$$\begin{aligned}
\beta(z) &= \frac{\lambda}{1-\rho_a}\left(2c(a)\overline{F}(x) + \int_0^\infty \beta(u)\overline{F}(a+z+u)\,du\right. \\
&\qquad \left. + \int_0^z \beta(u)\overline{F}(a+z-u)\,du\right), \qquad (27)
\end{aligned}$$

where

$$c(a) = \frac{\partial}{\partial a}\overline{T}^{\text{FB}}(a) = \frac{1+2\lambda\overline{T}^{\text{FB}}(a)\overline{F}(a)}{1-\rho_a} > 0.$$

Now we will show that $\beta(z) \ge 0$ for all $z > 0$. For that, define

$$\beta^* = \inf_{z\ge 0}\beta(z) > -\infty.$$

By (27), we have, for all $z \geq 0$,

$$\begin{aligned}\beta(z) &\geq \frac{\lambda\beta^*}{1-\rho_a}\left(\int_0^\infty \overline{F}(a+z+u)\,du + \int_0^z \overline{F}(a+z-u)\,du\right)\\ &= \frac{\lambda\beta^*}{1-\rho_a}(E[S]-E[S\wedge a])\\ &= \beta^*\frac{\rho-\rho_a}{1-\rho_a}.\end{aligned}$$

Thus,

$$\beta^* \geq \beta^* \frac{\rho-\rho_a}{1-\rho_a},$$

implying that $\beta^* \geq 0$, since

$$0 \leq \frac{\rho-\rho_a}{1-\rho_a} < 1.$$

As an immediate consequence, we get that, for all $x > a$,

$$(g_x)'(a) = \beta(x-a) \geq 0.$$

Since this is true for any a we finally deduce that $g_x(a) = (\overline{T}^{2\mathrm{PS}(a)})'(x)$ is an increasing function of the level threshold a. Thus, the claim is true. □

Let us now recall Proposition 7 and then prove it.

Proposition 7 *Let $N \geq 1$, $n \in \{1,\ldots,N\}$, $\pi \in \mathrm{MLPS}$ with thresholds $\{a_1,\ldots,a_N\}$, and $\pi' \in \mathrm{MLPS}$ with thresholds $\{a_1,\ldots,a_{n-1},a_{n+1},\ldots,a_N\}$ such that $D_n^{\pi'} = D_n^\pi = D_{n+1}^\pi = \mathrm{PS}$ and*

$$D_m^{\pi'} = \begin{cases} D_m^\pi & m = 1,\ldots,n-1,\\ D_{m+1}^\pi & m = n+1,\ldots,N.\end{cases}$$

Then, $\overline{U}_x^\pi \leq \overline{U}_x^{\pi'}$ for all $x \geq 0$.

Proof. 1° Assume first that $n = N$, i.e., the highest level is splitted. By Propositions 2 and 10, we deduce that $\overline{U}_x^\pi \leq \overline{U}_x^{\pi'}$ for all $x \leq a_N$. In particular,

$$\overline{U}_{a_N}^\pi \leq \overline{U}_{a_N}^{\pi'}. \tag{28}$$

Regarding the remaining values, we may, instead of π, consider the two-level discipline $2\mathrm{PS}(a_N)$, since, for all $x \geq a_N$,

$$\overline{U}_x^\pi = \overline{U}_x^{2\mathrm{PS}(a_N)},$$

again by Proposition 2. Similarly, instead of π', we may consider the two-level discipline $2\mathrm{PS}(a_{N-1})$, since, for all $x \geq a_{N-1}$,

$$\overline{U}_x^{\pi'} = \overline{U}_x^{2\mathrm{PS}(a_{N-1})}.$$

As $x \to \infty$, we get the ordinary mean unfinished work,

$$\overline{U}_\infty^\pi = \overline{U}_\infty^{2\mathrm{PS}(a_N)} = \overline{U}_\infty^{2\mathrm{PS}(a_{N-1})} = \overline{U}_\infty^{\pi'}. \tag{29}$$

Proposition 12 tells us that, for all $x > a_N$,

$$(\overline{T}^{2\mathrm{PS}(a_{N-1})})'(x) \le (\overline{T}^{2\mathrm{PS}(a_N)})'(x).$$

Thus, the difference

$$\overline{T}^{2\mathrm{PS}(a_N)}(x) - \overline{T}^{2\mathrm{PS}(a_{N-1})}(x)$$

is non-decreasing for all $x \ge a_N$. By combining this with (28) and (29), we conclude from Lemma 1 that, for all $x \ge a_N$,

$$\overline{U}_x^\pi = \overline{U}_x^{2\mathrm{PS}(a_N)} \le \overline{U}_x^{2\mathrm{PS}(a_{N-1})} = \overline{U}_x^{\pi'}.$$

So, the claim is true for $n = N$.

2° Assume then that $n < N$. By Proposition 2, we deduce that, for all $x \le a_{n-1}$ and $x \ge a_{n+1}$,

$$\overline{U}_x^\pi = \overline{U}_x^{\pi'}.$$

So, it remains to consider values $a_{n-1} \le x \le a_{n+1}$.

Let $a_{n-1} \le x \le a_{n+1}$. Recall from Subsection 4.1 that $\overline{U}_x^\pi(S \wedge b)$ denotes the mean unfinished truncated work in a system where the original service times $S_1, S_2, \ldots$ are replaced by their truncated versions $S_1 \wedge b, S_2 \wedge b, \ldots$. By Proposition 11,

$$\overline{U}_x^\pi = \overline{U}_x^\pi(S \wedge a_{n+1}), \quad \overline{U}_x^{\pi'} = \overline{U}_x^{\pi'}(S \wedge a_{n+1}).$$

Let then $\pi'' \in 3\mathrm{PS}$ with thresholds $\{a_{n-1}, a_n\}$ and $\pi''' \in 2\mathrm{PS}$ with threshold a_{n-1}. In fact, for $n = 1$, $\pi'' \in 2\mathrm{PS}$ with threshold a_1 and $\pi''' = \mathrm{PS}$. Note that π'' is generated, in any case, from π''' by splitting the highest level. By applying Proposition 2 to the system with truncated service times, we get

$$\begin{aligned} \overline{U}_x^\pi(S \wedge a_{n+1}) &= \overline{U}_x^{\pi''}(S \wedge a_{n+1}), \\ \overline{U}_x^{\pi'}(S \wedge a_{n+1}) &= \overline{U}_x^{\pi'''}(S \wedge a_{n+1}). \end{aligned}$$

Finally, by applying the result of 1° to the system with truncated service times, we deduce that

$$\overline{U}_x^{\pi''}(S \wedge a_{n+1}) \le \overline{U}_x^{\pi'''}(S \wedge a_{n+1}).$$

This completes the proof of 2°, and, thus also the whole proof of Proposition 7. □

5 Numerical examples

In this section we present illustrative numerical examples concerning various MLPS disciplines in the context of M/G/1 queues with hyperexponential and Pareto service time distributions. We note that both distributions belong to the class DHR.

Hyperexponential service times

In the first example we have used a hyperexponential service time distribution with tail distribution function

$$\overline{F}(x) = pe^{-\mu_1 x} + (1-p)e^{-\mu_2 x},$$

where $\mu_1 = 1.0$, $\mu_2 = 0.1$, and $p = 0.8$. For this distribution, $E[S] = 2.8$. The arrival rate is $\lambda = 0.3$, implying load $\rho = 0.84$ and mean unfinished work $\overline{U}_\infty = 39.0$.

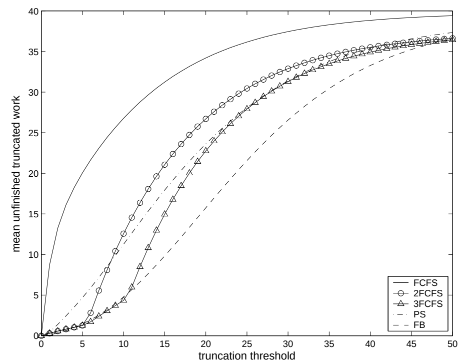

Figure 1: Hyperexponential service times: Mean unfinished truncated work $\overline{U}_x^\pi$ as a function of the truncation threshold x for disciplines FCFS, 2FCFS(5), 3FCFS(5, 10), PS, and FB.

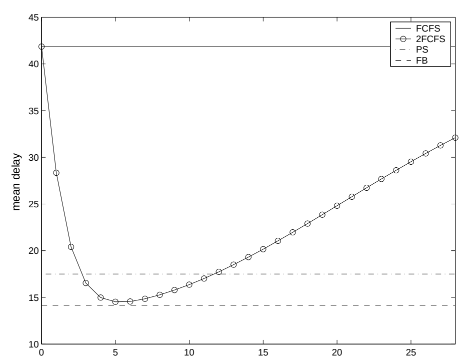

Figure 2: Hyperexponential service times: Mean delay $\overline{T}^\pi$ as a function of the level threshold a for discipline 2FCFS(a). The three horizontal lines correspond to the mean delay of disciplines FCFS, PS and FB.

In Figure 1, we have depicted the mean unfinished truncated work $\overline{U}_x^\pi$ as a function of the truncation threshold x for MLPS disciplines FCFS, 2FCFS(5), 3FCFS(5, 10), PS, and FB. As claimed in Proposition 6, the mean unfinished truncated work is decreasing uniformly when the number of levels is increased from one (FCFS) to two (2FCFS(5)), to three (3FCFS(5, 10)), and to infinity (FB).

In Figure 2, we have depicted the mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for discipline 2FCFS(a). As claimed in Theorem 3, the mean delay for any 2FCFS(a) discipline is less than that of FCFS. It is also interesting to observe that, for a quite wide range of thresholds a, the two-level disciplines 2FCFS(a) give even lower mean delay than PS. In fact, the optimal two-level discipline is almost as good as FB. Thus, if the first threshold is chosen optimally, the gain that could be obtained by adding more levels is negligible.

Pareto service times

The rest of our examples are based on the Pareto service time distribution with tail distribution function

$$\overline{F}(x) = \left(\frac{1}{1+cx}\right)^{\alpha}.$$

We vary parameters c and α while keeping the mean service time fixed $E[S] = 30.0$.

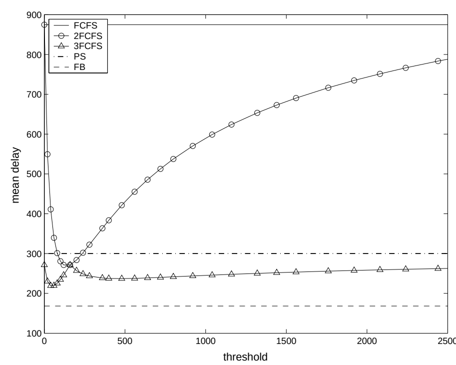

Figure 3: Pareto service times: Mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2FCFS(a) and 3FCFS($\min\{a, a^*\}, \max\{a, a^*\}$) with $\alpha = 2.2$. The three horizontal lines correspond to the mean delay of disciplines FCFS, PS, and FB.

First we consider the case $c = 1/36$ and $\alpha = 2.2$. The arrival rate is $\lambda = 0.03$, implying load $\rho = 0.9$. In Figure 3, we have depicted the mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2FCFS(a) and 3FCFS($\min\{a, a^*\}, \max\{a, a^*\}$). The threshold value a^* used for the three-level disciplines is chosen to be the optimal one among the two-level disciplines 2FCFS(a). Note that again the optimal two-level discipline 2FCFS(a^*) gives even lower mean delay than PS. Having a third level gives some additional gain.

In the following example, we have a heavy tailed distribution with parameters $c = 1/24$ and $\alpha = 1.8$. In this case $E[S^2] = \infty$ so that the mean delay of the FCFS discipline is unbounded. The arrival rate is $\lambda = 0.03$, implying load $\rho = 0.9$. In Figure 4, we have depicted the mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2PS(a) and 3PS($\min\{a, a^*\}, \max\{a, a^*\}$). The threshold value a^* used

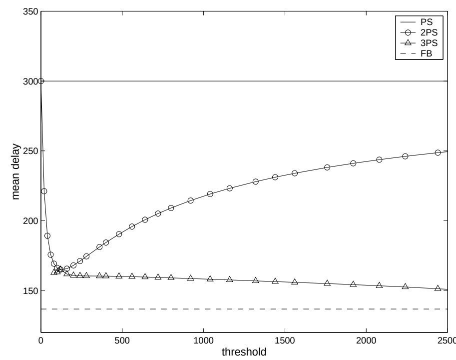

Figure 4: Pareto service times: Mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2PS(a) and 3PS($\min\{a, a^*\}, \max\{a, a^*\}$) with $\alpha = 1.8$. The two horizontal lines correspond to the mean delay of disciplines PS and FB.

for the three-level disciplines is chosen to be the optimal one among the two-level disciplines 2PS(a).

Next we take an even heavier tailed distribution with parameters $c = 1/6$ and $\alpha = 1.2$. Thus, again $E[S^2] = \infty$. The arrival rate is now $\lambda = 0.02$, implying load $\rho = 0.6$. In Figure 5, we have again depicted the mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2PS(a) and 3PS($\min\{a, a^*\}, \max\{a, a^*\}$). In this case, the performance of two-level or three-level policies is rather insensitive with respect to the chosen thresholds.

In the final example we have used parameters $c = 1/45$ and $\alpha = 2.5$ with $E[S^2] < \infty$. The arrival rate is $\lambda = 0.03$, implying load $\rho = 0.9$. In Figure 6, we have depicted the mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2FCFS(a) and 2PS(a). As claimed in Theorem 2, the mean delay for any 2PS(a) discipline is less than that of the corresponding 2FCFS(a) discipline. The shapes of the two curves are similar and the optimal level thresholds are not far away from each other. In addition, we observe that by choosing the level threshold optimally the 2FCFS discipline gives a remarkable gain compared to the FCFS discipline, performing even better than PS. The optimal 2PS discipline is almost as good as FB.

In conclusion, a small number of levels seems to give good results if at least one of the thresholds is reasonably chosen with respect to the prevailing traffic conditions. Thus, if these traffic conditions are known, two levels seem to be enough. On the other hand, if they are only partially known or they are rapidly changing, it would be reasonable to use an MLPS policy with a couple of levels for which the level thresholds are chosen from different magnitudes, for example, $a_n = a^n$ for a suitable constant a.

6 Conclusions

This paper compares the mean delay among MLPS disciplines. The MLPS disciplines form a dense subset within the set of non-anticipating scheduling disciplines. Previous

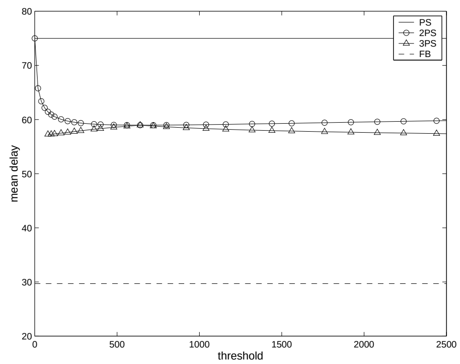

Figure 5: Pareto service times: Mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2PS(a) and 3PS(min$\{a, a^*\}$, max$\{a, a^*\}$) with $\alpha = 1.2$. The two horizontal lines correspond to the mean delay of disciplines PS and FB.

analysis considered FB and PS as internal disciplines and showed that such MLPS disciplines are better than PS with respect to the mean delay whenever the hazard rate of the service time distribution is decreasing. In this article we allow FCFS as an internal discipline as well. Under the same DHR condition as above, we notably show that given any MLPS discipline, the mean delay is reduced if a level is added by splitting an existing one. As the number of levels increases, the MLPS discipline emulates closer and closer the behavior of an FB discipline, which is known to be optimal. By numerical analysis we have evaluated the mean delay of MLPS discipline in the presence of distributions of practical interest as Pareto and hyperexponential. Our results indicate that if the level threshold is chosen appropriately, the mean delay of a two-level MLPS can be close to that of FB. This result is important in view of recent work that proposes to provide differential treatment to flows on the Internet based in just two classes: mice and elephants.

An important problem to be studied in the future is the performance of MLPS disciplines when there are more than just one queue. In agreement with the result by Bonald and Massoulié [18], we expect that the region of stability will be reduced with respect to the non-priority case. A more thorough study of the advantages/disadvantages is, however, required in order to understand better the deployment and scalability on the Internet of such priority mechanisms.

A Truncation Lemma

In this appendix we present a lemma that is needed for the proof of Proposition 9 presented in Section 3.

Lemma 2 *Assume that there are two sets of real numbers,* $(a_1, \ldots, a_N)$ *and* $(b_1, \ldots, b_N)$*, and an index* $1 \le m \le N$ *that satisfy the following conditions:*

(a) $\sum_{i=1}^{N} a_i = \sum_{i=1}^{N} b_i$,

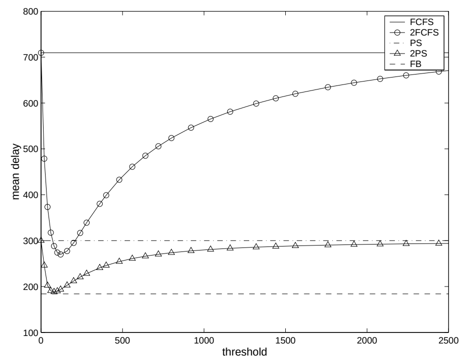

Figure 6: Pareto service times: Mean delay $\overline{T}^{\pi}$ as a function of the level threshold a for disciplines 2FCFS(a) and 2PS(a) with $\alpha = 2.5$. The three horizontal lines correspond to the mean delay of disciplines FCFS, PS, and FB.

(b) $a_i \le b_i$ *for all* $i \in \{1, \ldots, m\}$,

(c) $a_i \ge b_i$ *for all* $i \in \{m+1, \ldots, N\}$,

(d) $a_i \ge a_j$ *for all* $i \in \{1, \ldots, m\}$ *and* $j \in \{i+1, \ldots, N\}$.

Then, for any $x \ge 0$,

$$\sum_{i=1}^{N} (a_i \wedge x) \ge \sum_{i=1}^{N} (b_i \wedge x).$$

Proof. Note first that, for any $n \le m$,

$$\begin{aligned} \sum_{i=n+1}^{N} a_i &= \sum_{i=1}^{N} a_i - \sum_{i=1}^{n} a_i \\ &\overset{\text{(a)}}{=} \sum_{i=1}^{N} b_i - \sum_{i=1}^{n} a_i \\ &\overset{\text{(b)}}{\ge} \sum_{i=1}^{N} b_i - \sum_{i=1}^{n} b_i = \sum_{i=n+1}^{N} b_i. \end{aligned} \tag{30}$$

Denote $A = \max\{a_{m+1}, \ldots, a_N\}$, and consider separately cases $x > A$ and $x \le A$.

1° Let $x > A$ and define

$$p = \begin{cases} 0 & \text{if } a_1 < x, \\ \max\{i \in \{1, \ldots, m\} : a_i \ge x\}, & \text{otherwise.} \end{cases}$$

By the definition above and assumptions (b) and (d), we deduce that $b_i \ge a_i \ge x$ for all $i \in \{1, \ldots, p\}$, while $a_i < x$ for all $i \in \{p+1, \ldots, m\}$. Furthermore $a_i \le A < x$ for all $i \in \{m+1, \ldots, N\}$. Thus,

$$\begin{aligned}
\sum_{i=1}^{N}(a_i \wedge x) &= \sum_{i=1}^{p} x + \sum_{i=p+1}^{N} a_i \\
&\stackrel{(30)}{\geq} \sum_{i=1}^{p} x + \sum_{i=p+1}^{N} b_i \\
&\geq \sum_{i=1}^{N}(b_i \wedge x).
\end{aligned}$$

2° Let then $x \leq A$. By assumption (d), $a_i \geq A \geq x$ for all $i \in \{1, \ldots, m\}$. Thus,

$$\begin{aligned}
\sum_{i=1}^{N}(a_i \wedge x) &= \sum_{i=1}^{m} x + \sum_{i=m+1}^{N} (a_i \wedge x) \\
&\stackrel{(c)}{\geq} \sum_{i=1}^{m} x + \sum_{i=m+1}^{N} (b_i \wedge x) \\
&\geq \sum_{i=1}^{N}(b_i \wedge x).
\end{aligned}$$

This completes the proof. □

B Batch arrival PS queue

Consider an $M^X/G/1$ queue with batch arrivals and PS discipline. This appendix gives a summary of some important properties of such a queueing system, which is closely related to ordinary M/G/1 queues with $(N+1)$PS disciplines, as explained in [1].

Let λ denote the rate at which batches arrive and Y the size of an arriving batch with

$$\bar{a} = E[Y], \qquad b = \frac{E[Y^2]}{E[Y]} - 1.$$

Note that b refers to the mean number of additional customers in an arriving batch where there is at least one customer. Furthermore, let $B(t) = P\{S \leq x\}$ denote the service time distribution, and $\overline{B}(t) = 1 - B(t)$ the corresponding tail distribution. Assume that the system is stable, i.e., $\rho = \lambda E[Y]E[S] < 1$, where ρ denotes the load.

Let then $\alpha(t)$ denote the conditional mean delay for a customer with service requirement t. As found in [19], we have the following integral equation for the derivative $\alpha'(t) = d\alpha(t)/dt$, $t \geq 0$:

$$\begin{aligned}
\alpha'(t) &= 1 + b\overline{B}(t) + \lambda\bar{a}\int_0^\infty \alpha'(u)\overline{B}(t+u)\,du \\
&\quad + \lambda\bar{a}\int_0^t \alpha'(u)\overline{B}(t-u)\,du. \qquad (31)
\end{aligned}$$

The term $\alpha'(t)$ is proportional to the average time required for a job that has already attained t units of service, to get served an extra infinitesimal amount of service. Hence, in a Processor-Sharing system, $\alpha'(t)$ is equal to the average number of jobs present in the system when t units have been served. The first term on the right hand side accounts for the presence of the tagged job. The second term accounts for those jobs that arrived in the same batch together with the tagged job, and that are still present in the system. The third term accounts for those jobs that were present in the system upon arrival of the tagged job, and that are still present in the system when the tagged job has obtained t units of service. Finally, the fourth term accounts jobs that have arrived while the tagged job has been in the system.

References

[1] L. Kleinrock. *Queueing Systems, Volume II: Computer Applications.* John Wiley & Sons, New York, 1976.

[2] L. Guo and I. Matta. Differentiated control of web traffic: A numerical analysis. In *SPIE ITCOM'2002*, Boston, MA, 2002.

[3] H. Feng and V. Misra. Mixed scheduling disciplines for network flows. *ACM SIGMETRICS Performance Evaluation Review*, 31:36–39, 2003.

[4] K. Avrachenkov, U. Ayesta, P. Brown, and E. Nyberg. Differentiation between short and long TCP flows: predictability of the response time. In *IEEE Infocom 2004*, Hong Kong, 2004.

[5] I. Rai, G. Urvoy-Keller, M. Vernon, and E. Biersack. Performance analysis of LAS-based scheduling disciplines in a packet switched network. In *ACM SIGMETRICS/PERFORMANCE 2004*, pages 106–117, New York, NY, 2004.

[6] S. Yashkov. Processor-sharing queues: Some progress in analysis. *Queueing Systems*, 2:1–17, 1987.

[7] R. Righter and J. Shanthikumar. Scheduling multiclass single server queueing systems to stochastically maximize the number of successful departures. *Probability in the Engineering and Informational Sciences*, 3:323–333, 1989.

[8] D. Heyman, T. Lakshman, and A. Neidhardt. A new method for analysing feedback-based protocols with applications to engineering web traffic over the Internet. In *ACM SIGMETRICS 1997*, pages 24–38, Seattle, WA, 1997.

[9] M. Nabe, M. Murata, and H. Miyahara. Analysis and modelling of world wide web traffic for capacity dimensioning of internet access lines. *Performance Evaluation*, 34(4):249–271, 1998.

[10] L. Massoulié and J. Roberts. Bandwidth sharing and admission control for elastic traffic. *Telecommunication Systems*, 15(1-2):185–201, 2000.

[11] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie, and J. Roberts. Statistical bandwidth sharing: A study of congestion at flow level. In *ACM SIGCOMM 2001*, pages 111–122, San Diego, CA, 2001.

[12] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. In *ACM SIGMETRICS 1996*, pages 160–169, Philadelphia, PA, 1996.

[13] A. Feldmann and W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. In *IEEE Infocom 1997*, pages 1096–1104, Kobe, Japan, 1997.

[14] S. Aalto, U. Ayesta, and E. Nyberg-Oksanen. Two-level processor-sharing scheduling disciplines: mean delay analysis. In *ACM SIGMETRICS/PERFORMANCE 2004*, pages 97–105, New York, NY, 2004.

[15] S. Aalto, U. Ayesta, and E. Nyberg-Oksanen. M/G/1/MLPS compared to M/G/1/PS. Research Report RR-5219, INRIA, 2004. To appear in *Operations Research Letters*.

[16] S. Halfin and W. Whitt. An extremal property of the FIFO discipline via an ordinal version of $L = \lambda W$. *Stochastic Models*, 5:515–529, 1989.

[17] R. Righter, J. Shanthikumar, and G. Yamazaki. On extremal service disciplines in single-stage queueing system. *Journal of Applied Probability*, 27:409–416, 1990.

[18] T. Bonald and L. Massoulié. Impact of fairness on internet performance. In *ACM SIGMETRICS 2001*, pages 82–91, Cambridge, MA, 2001.

[19] L. Kleinrock, R. Muntz, and E. Rodemich. The processor-sharing queueing model for time-shared systems with bulk arrivals. *Networks Journal*, 1:1–13, 1971.